# Imprinting detection by extending a regression-based QTL analysis method

Statistical Genetics Journal Club

Jingyuan Yang

Department of Statistics
The Ohio State University

Sep. 26th 2007

# Linkage Analysis

- **Haseman and Elston (1970)** For sibling pair $s$, at marker $m$: $(X_{s1} - X_{s2})^2 = b_0 + b_1\hat{\pi}_{sm} + \epsilon_s$, $E[b_1] = -2(1-2c)^2\sigma_g^2$
- **Amos (1994)** $R$ pedigrees are sampled, each having $n_r$ members. For pedigree $r$, assume: $\mathbf{X_r} \sim MVN_{n_r}(\boldsymbol{\mu}_\mathbf{r}, \boldsymbol{\Sigma}_\mathbf{r})$ Where, $\boldsymbol{\Sigma}_\mathbf{r}$ involves parameters $\sigma_a^2$, $\sigma_d^2$, $\sigma_G^2$, and $c$, etc.
- **Sham et.al (2002)** $\hat{\Pi}_C = \Sigma'_{Y\hat{\Pi}}\Sigma_Y^{-1}\mathbf{Y}_C + \epsilon$ Claimed to combine the simplicity and robustness of regression-based methods and the generality and greater power of VC models.

# Imprinting (Parent-of-origin effect)

An epigenetic alteration of genes in which primarily the maternally or paternally inherited allele is expressed.

Introduces functional (expressional) inequility between 2 parental alleles of a gene.

Causes deviations from Mendelian law, therefore, conventional linkage analysis methods are not appropriate and require modification.

## Sham's method

Data: Squared sums **S**, squared differences **d** (Let **Y** denote $[S, d]$')and estimated IBD sharing $\hat{\Pi}$

Compute:

- Covariance Matrices of **S** and **d**
- Covariance Matrix of $\hat{\Pi}$
- Covariance between **S**, **d** and $\hat{\Pi}$

From these, $\Sigma_Y$ and $\Sigma_{Y\hat{\Pi}}$ can be obtained.

Then regress $\hat{\Pi}_C$ on $\mathbf{Y}_C$: $\hat{\Pi}_C = \Sigma'_{Y\hat{\Pi}} \Sigma_Y^{-1} \mathbf{Y}_C + \epsilon$

# Sham's method - continued

Matrix $\Sigma'_{Y\hat{\Pi}}$ is proved to be equal to $Q\Sigma_{\hat{\Pi}}H$, where

$Q$ is the phenotypic variance explained by the additive effects of the QTL,

$H$ is a constant matrix with all elements being equal to $\pm 2$ or 0.

$E(\hat{\Pi}_C) = Q\Sigma_{\hat{\Pi}}H\Sigma_Y^{-1}\mathbf{Y}_C = Q\Sigma_{\hat{\Pi}}B$ (denote $H\Sigma_Y^{-1}\mathbf{Y}_C$ by $B$)

Then the estimate of $Q$ is $B'\hat{\Pi}_C/B'\Sigma_{\hat{\Pi}}B$, with sampling variance $1/B'\Sigma_{\hat{\Pi}}B$.

Test statistic $T = \hat{Q}\sum[B'\hat{\Pi}_C] = \hat{Q}^2\sum[B'\Sigma_{\hat{\Pi}}B]$

# Test for Imprinting

Use overall $\hat{\pi}$ and parent-specific estimated IBD sharings $\hat{\pi}_p$ and $\hat{\pi}_m$, we can get three test statistics $T$, $T_p$ and $T_m$, as well as three estimates of additive variance component $\hat{Q}$, $\hat{Q}_p$ and $\hat{Q}_m$.

To test imprinting, they proposed:

- Compare $T$, $T_p$ and $T_m$, or
- Compute a new test statistic: $I = \frac{\hat{Q}_p - \hat{Q}_m}{\sqrt{Var(\hat{Q}_p) + Var(\hat{Q}_m)}}$

  To perform the I test in practice , one can:
  - Get empirical critical values from data simulated under null hypothesis, or
  - Obtain the reference distribution for I by using a permutation technique.

## Computation of Parent-Specific IBD

Let $x_0$ and $y_0$ be the marker phenotypes of parents, and let $x_i$, i=1,...,p, be the marker phenotypes of $p$ sibs.

Let $g_{ab}(x)$=P(observing phenotype $x$|genotype is $ab$).

Population frequency of genotype $ab$ is denoted by $\phi_{ab}$.

Phenoset: Set of all genotypes that could gie rise to the observed phenotype.

Let $F$ and $M$ be the phenoset of father and mother, respectively.

# Computation of Parent-Specific IBD - Continued

Let $rs$ be an element of $F$, $vw$ be an element of $M$.

L(family data)= $\sum_{rs} \phi_{rs} g_{rs}(x_0) \sum_{vw} \phi_{vw} g_{vw}(y_0) \prod_{j=1}^{p} \frac{1}{4}[g_{rv}(x_j) + g_{rw}(x_j) + g_{sw}(x_j) + g_{sv}(x_j)]$

L(family data and sibs 1 and 2 having one allele IBD from the father)= $\sum_{rs} \phi_{rs} g_{rs}(x_0) \sum_{vw} \phi_{vw} g_{vw}(y_0) \prod_{j=3}^{p} \frac{1}{4}[g_{rv}(x_j) + g_{rw}(x_j) + g_{sw}(x_j) + g_{sv}(x_j)] \times \frac{1}{16}[g_{rv}(x_1)g_{rw}(x_2) + g_{rw}(x_1)g_{rv}(x_2) + g_{sv}(x_1)g_{sw}(x_2) + g_{sw}(x_1)g_{sv}(x_2) + g_{rv}(x_1)g_{rv}(x_2) + g_{rw}(x_1)g_{rw}(x_2) + g_{sw}(x_1)g_{sw}(x_2) + g_{sv}(x_1)g_{sv}(x_2)]$

$\hat{\pi}_{p12} = P(\pi_{p12} = 1|\text{family data})=$
L(family data and sibs 1 and 2 having one allele IBD from the father)/L(family data)